

Linear dynamics for the state vector of Markov chain functions

James Ledoux*

01 September 2004

Abstract

Let $(\varphi(X_n))_n$ be a function of a finite-state Markov chain $(X_n)_n$. In this note, we investigate under which conditions the random variable $\varphi(X_n)$ have the same distribution as Y_n (for every n), where $(Y_n)_n$ is a Markov chain with fixed transition probability matrix. In other words, for a deterministic function φ , we investigate the conditions under which $(X_n)_n$ is *weakly lumpable for the state vector*. We show that the set of all probability distributions of X_0 such that $(X_n)_n$ is weakly lumpable for the state vector can be finitely generated. The connections between our definition of lumpability and usual one's, as the proportional dynamics property, are discussed.

KEYWORDS: Lumpability, Proportional dynamics, Rogers-Pitman matrix

AMS: 60J10

1 Introduction

Let us consider a discrete-time S -valued homogeneous Markov chain (HMC), $(X_n)_n$ with $S = \{1, \dots, N\}$. Its transition probability matrix (tpm) is denoted by P . Such a stochastic process is fully specified by the probability distribution $\text{distr}(X_0)$ of the random variable X_0 and the matrix P . A map φ from $S = \{1, \dots, N\}$ into $\hat{S} = \{1, \dots, \hat{N}\}$ such that $\varphi(S) = \hat{S}$, (with $\hat{N} < N$), is called a *lumping map*. We see that a lumping map φ is specified by the $N \times \hat{N}$ -matrix V_φ

$$x \in S, y \in \hat{S}, \quad V_\varphi(x, y) := 1 \text{ if } x \in \varphi^{-1}(y) \text{ and } 0 \text{ otherwise.}$$

*IRMAR UMR 6625 & INSA de Rennes, 20 avenue des Buttes de Coësmes, 35708 Rennes cedex 7, France

We are interested in the lumped process $(\varphi(X_n))_n$. Specifically, we consider the following linear dynamic system

$$\forall n \geq 0, \quad \begin{cases} \alpha_{n+1} &= \alpha_n P \\ \hat{\alpha}_n &= \alpha_n V_\varphi \end{cases}$$

where α_n and $\hat{\alpha}_n$ are stochastic vectors corresponding to the probability distributions of the random variables X_n and $\varphi(X_n)$ respectively. When $(\varphi(X_n))_n$ is a Markov chain for a fixed $\text{distr}(X_0)$, $(X_n)_n$ is said to be *weakly lumpable* [9]. The corresponding basic properties are stated in [9], [17] and [10]. In this paper, we are only interested in a concept of *weak lumpability for the state vector*. That means, there exists a $\hat{N} \times \hat{N}$ stochastic matrix \hat{P} such that, for a fixed $\text{distr}(X_0)$, the process $(\varphi(X_n))_n$ has the same one-dimensional distributions than the \hat{S} -valued Markov chain specified by the tpm \hat{P} and the probability distribution $\text{distr}(\varphi(X_0))$. In other words, the state probabilities vectors $(\hat{\alpha}_n)_{n \geq 0}$, associated with the lumped process $(\varphi(X_n))_n$, satisfy the difference equation

$$\begin{cases} \hat{\alpha}_{n+1} &= \hat{\alpha}_n \hat{P} \\ \hat{\alpha}_0 &= \alpha_0 V_\varphi. \end{cases} \quad n \geq 0$$

When it holds for any $\text{distr}(X_0)$, $(X_n)_n$ is said to be *strongly lumpable*. In this last case, we recover the concept of lumpability of a general linear dynamic system as discussed in many papers and books (see e.g. [1], [5]).

Our work originates in the papers [3] and [13, 14] which attempt to simplify the computation of the transient characteristics of a Markov model using lumping. Our concept of lumpability is clearly relevant when assessing the probability distributions $\text{distr}(X_n), n \geq 1$. It may be also of some value when we are concerned with the computation of a measure which only depends on the occupation probabilities of a subset of states by $(X_n)_n$. Let us briefly illustrate our purpose on the following example.

Example 1.1

We consider the following Markov chain. It is derived from the reliability model described Page 10. Its tpm is

$$P = \left(\begin{array}{ccc|ccc} 1/8 & 1/8 & 2/8 & 2/8 & 1/8 & 1/8 \\ 3/16 & 1/16 & 4/16 & 4/16 & 3/16 & 1/16 \\ 2/8 & 2/8 & 2/8 & 0 & 1/8 & 1/8 \\ \hline 0 & 0 & 2/8 & 4/8 & 1/8 & 1/8 \\ 1/8 & 1/8 & 2/8 & 2/8 & 1/8 & 1/8 \\ 1/16 & 3/16 & 4/16 & 4/16 & 3/16 & 1/16 \end{array} \right). \quad (1.1)$$

The lumping map is defined by $\varphi(1) = \varphi(2) = \varphi(3) = 1$, $\varphi(4) = \varphi(5) = \varphi(6) = 2$. The corresponding matrix V_φ is

$$V_\varphi = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix}^\top.$$

Let us consider

$$\hat{P} = \begin{pmatrix} 5/8 & 3/8 \\ 3/8 & 5/8 \end{pmatrix} \quad (1.2)$$

as matrix \hat{P} .

If $(\varphi(X_n))$ was a Markov chain with tpm \hat{P} for the probability distribution $\text{distr}(X_0)$, then **every** measure that only depends on the occupation of the subset of states $\{1, 2, 3\}$ and $\{4, 5, 6\}$ by $(X_n)_n$, could be computed from the $\{1, 2\}$ -valued Markov model $(\text{distr}(\varphi(X_0)), \hat{P})$. For instance, let us consider the probability mass function of the cumulative occupation time of the subset $\{1, 2, 3\}$ or $\{4, 5, 6\}$

$$\forall n \geq 1, \quad C_{i,n} := \sum_{k=1}^n \mathbf{1}_{\{X_k \in \varphi^{-1}(i)\}} = \sum_{k=1}^n \mathbf{1}_{\{\varphi(X_k)=i\}}, \quad i = 1, 2.$$

A closed-form of the distribution $\mathbb{P}\{C_{i,n} = \cdot\}$, $i = 1, 2$ is provided in [18]. This is based on difference equations which are very costly in terms of matrix computations. When weak lumpability holds, we can compute this probability mass function from the two-states Markov chain $(\text{distr}(\varphi(X_0)), \hat{P})$ by standard scalar equations. Unfortunately, we shall show that $(\varphi(X_n))$ is not a Markov chain, for all initial probability distribution $\text{distr}(X_0)$.

If we now assume that we are only interested in the expected cumulative occupation times

$$\begin{aligned} \forall n \geq 1, \quad (\mathbb{E}[C_{i,n}])_{i=1,2} &= \left(\sum_{k=1}^n \mathbb{P}\{X_k \in \varphi^{-1}(i)\} \right)_{i=1,2} \\ &= \left(\sum_{k=1}^n \mathbb{P}\{\varphi(X_k) = i\} \right)_{i=1,2} = \sum_{k=1}^n \alpha P^k V_\varphi \end{aligned}$$

where α is the stochastic vector corresponding to the probability distribution $\text{distr}(X_0)$. We set $\alpha = (1/6, 1/3, 1/6, 0, 1/6, 1/6)$. It is an easy computation to verify that

$$\forall n \geq 1, \quad (\mathbb{E}[C_{i,n}])_{i=1,2} = (2/3, 1/3) \sum_{k=1}^n \hat{P}^k,$$

with \hat{P} the 2×2 -matrix given in (1.2). In fact, we shall obtain that the above equality still holds when $\text{distr}(X_0)$ is any convex combination of the stochastic vectors

$$\begin{pmatrix} 1/2 & 0 & 1/2 & 0 & 0 & 0 \\ 1/2 & 0 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1/2 & 0 & 1/2 \end{pmatrix} \quad \begin{pmatrix} 0 & 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 0 & 0 & 1/2 \end{pmatrix} \quad \begin{pmatrix} 0 & 0 & 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 & 0 \\ 1/2 & 0 & 0 & 0 & 0 & 1/2 \end{pmatrix}.$$

In such cases, $(\text{distr}(X_0), P)$ is said to be weakly lumpable for the state vector. For such a $\text{distr}(X_0)$, the computation of transient measures related to (and only to) the probability

function mass of random variables $\mathbf{1}_{\{X_n \in \varphi^{-1}(i)\}}$ $i = 1, 2$, may be carried out from the two-states Markov chain $(\text{distr}(X_0)V_\varphi, \hat{P})$. Instances of such measures are the point availability in reliability, the expected accumulated reward in performance evaluation, or the output process of some discrete-time queues.

The aim of this note is to establish basic properties associated with the notion of lumpability introduced above. We also document its relationship with different concepts of lumpability used in [3, 13, 14], in particular with the so-called proportional dynamics property.

The article is organized as follows. In Section 2, we analyze the set all probability distributions $\text{distr}(X_0)$ such that $(X_n)_n$ is weakly lumpable for the state vector. In particular, we show that this set can be finitely generated. In Section 3, we discuss a connection between our lumpability property and the Chapman-Kolmogorov condition for the lumped process $(\varphi(X_n))_n$. In Section 4, we investigate the connections between a lumpability condition used in [3], the proportional dynamics property as defined in [13, 14] and the lumpability property introduced above. We conclude in Section 5.

Basic notation

By convention, the vectors are row vectors. The i th component of a vector α is denoted by $\alpha(i)$. The column vectors are indicated by means of the transpose operator $(.)^\top$. The vector having all components equal to 1 (resp. 0) is merely denoted by $\mathbf{1}$ (resp. $\mathbf{0}$). The entry (i, j) of a matrix A is denoted by $A(i, j)$. The matrix \hat{I} denotes the $\hat{N} \times \hat{N}$ unity matrix.

An HMC specified by the tpm P and the initial probability distribution $\text{distr}(X_0)$ is denoted by $(\text{distr}(X_0), P)$.

A lumping map φ is assumed to be nondecreasing for notational convenience. If the stochastic vector α corresponds to $\text{distr}(X_0)$, then $\text{distr}(\varphi(X_0))$ is specified by the vector $(\mathbb{P}_\alpha\{\varphi(X_0) = y\})_{y \in \hat{S}} = \alpha V_\varphi$.

For any $y \in \hat{S}$, let Π_y be the $N \times N$ matrix defined by $\Pi_y(x, x) := 1$ if $x \in \varphi^{-1}(y)$ and $\Pi_y(x_1, x_2) := 0$ otherwise. This is the matrix of the orthogonal projector on the $\varphi^{-1}(y)$ -coordinates that we call the y -lumping projector.

For any nonnegative vector α such that $\alpha \Pi_y \mathbf{1}^\top > 0$, $\alpha^{(y)}$ denotes the vector $\alpha \Pi_y / \alpha \Pi_y \mathbf{1}^\top$. A $\hat{N} \times N$ nonnegative matrix U such that $U V_\varphi = \hat{I}$ is called a φ -block-diagonal matrix. For any nonnegative vector α such that $\alpha V_\varphi > \mathbf{0}$, the $\hat{N} \times N$ -matrix U_α defined by

$$y \in \hat{S}, x \in S \quad U_\alpha(y, x) := \alpha^{(y)}(x)$$

is φ -block-diagonal.

The definition and the basic properties of a cone used throughout this paper, are borrowed from [2]. For any $M \times N$ matrix G , $\text{Cone}(G)$ is the cone generated by the M rows of G .

2 Lumpability for the state probabilities vector

Let us define the concept of lumpability of a Markov chain for the state vector.

Definition 2.1 *Let \hat{P} be a $\hat{N} \times \hat{N}$ stochastic matrix. The HMC $(X_n)_n$ with tpm P is said to be weakly lumpable for the state vector with matrix \hat{P} , if there exists a probability distribution $\text{distr}(X_0)$ such that the process $(\varphi(X_n))_n$ has the same one-dimensional distributions than the Markov chain $(\text{distr}(\varphi(X_0)), \hat{P})$, that is, iff*

$$\forall n \geq 1, \quad \alpha P^n V_\varphi = \alpha V_\varphi \hat{P}^n \quad (2.1)$$

when $\text{distr}(X_0) = \alpha$. If Property (2.1) holds for any $\text{distr}(X_0)$, $(X_n)_n$ is said to be strongly lumpable (for the state vector).

Let us recall that the Markov chain $(X_n)_n$ has the standard weak lumpability property when $(\varphi(X_n))_n$ and $(\text{distr}(\varphi(X_0)), \hat{P})$ have the same finite dimensional distributions (see [10]). In other words, the stochastic processes $(\varphi(X_n))$ and $(\text{distr}(\varphi(X_0)), \hat{P})$ have the same distribution. When it is true for every $\text{distr}(X_0)$, we obtain the standard definition of strong lumpability. The usual weak lumpability property is clearly stronger than that corresponding to Definition 2.1. We emphasize that the two concepts are quite different in general but the difference has sometimes been overlooked in the literature (see e.g. [6, 20]). This may be explained by the following fact. The standard strong lumpability and the strong lumpability for the state vector are equivalent properties. This is proved below. Let us recall that $(X_n)_n$ with tpm P is strongly lumpable with \hat{P} (see [9]) iff

$$PV_\varphi = V_\varphi \hat{P}.$$

Proposition 2.1 *$(X_n)_n$ is strongly lumpable with the matrix \hat{P} iff $(X_n)_n$ is strongly lumpable for the state vector with \hat{P} . In this case, the matrix \hat{P} is $U_1 PV_\varphi$.*

Proof. We just have to justify the “if part”. It follows from Relation (2.1) with $n = 1$

$$\begin{aligned} \forall \alpha \geq \mathbf{0}, \quad \alpha PV_\varphi = \alpha V_\varphi \hat{P} &\iff \forall \alpha \geq \mathbf{0}, \quad \alpha [PV_\varphi - V_\varphi \hat{P}] = \mathbf{0} \\ &\iff \forall \alpha \in \mathbb{R}^N, \quad \alpha [PV_\varphi - V_\varphi \hat{P}] = \mathbf{0} \\ &\iff PV_\varphi = V_\varphi \hat{P}. \end{aligned}$$

Therefore, $(X_n)_n$ is strongly lumpable with the matrix \hat{P} . □

We are now interested in the set $\mathcal{D}_{\mathcal{M}}(\hat{P})$ of all $\text{distr}(X_0)$ such that $(X_n)_n$ is weakly lumpable for the state vector with matrix \hat{P} . $\mathcal{D}_{\mathcal{M}}(\hat{P})$ is clearly the set of the stochastic vectors belonging to

$$\mathcal{S}_{\mathcal{M}}(\hat{P}) := \{ \alpha \geq \mathbf{0} : \forall n \geq 1, \quad \alpha [P^n V_\varphi - V_\varphi \hat{P}^n] = \mathbf{0} \}.$$

If $\mathcal{S}_{\mathcal{M}}(\widehat{P}) \neq \{\mathbf{0}\}$ then $(X_n)_n$ is weakly lumpable for the state vector with \widehat{P} and

$$\mathcal{S}_{\mathcal{M}}(\widehat{P}) = \bigcap_{n \geq 1} \mathcal{S}_{\mathcal{M}}^{[n]} \quad (2.2)$$

where $\mathcal{S}_{\mathcal{M}}^{[n]}$ denotes the polyhedral cone

$$\mathcal{S}_{\mathcal{M}}^{[n]} := \{ \alpha \geq \mathbf{0} : \alpha[P^k V_{\varphi} - V_{\varphi} \widehat{P}^k] = \mathbf{0}, \text{ for } 1 \leq k \leq n \}. \quad (2.3)$$

2.1 Computation of $\mathcal{S}_{\mathcal{M}}(\widehat{P})$

The computation of $\mathcal{S}_{\mathcal{M}}(\widehat{P})$ may be carried out as follows. There is some similarity with the computation of $\mathcal{C}_{\mathcal{M}}(\widehat{P})$ in [17]. The comparison of these two sets is reported in Section 2.2.

Theorem 2.1 *The set $\mathcal{S}_{\mathcal{M}}(\widehat{P})$ may be finitely generated. We have*

$$\mathcal{S}_{\mathcal{M}}(\widehat{P}) = \mathcal{S}_{\mathcal{M}}^{[N+1]}.$$

Proof. For $n \geq 1$, we have $\mathcal{S}_{\mathcal{M}}^{[n+1]} \subseteq \mathcal{S}_{\mathcal{M}}^{[n]} \subseteq \mathcal{S}_{\mathcal{M}}^{[1]}$ from Definition (2.3) of the polyhedral cone $\mathcal{S}_{\mathcal{M}}^{[n]}$. Now, let us show that

$$\mathcal{S}_{\mathcal{M}}^{[n+1]} = \{ \alpha \in \mathcal{S}_{\mathcal{M}}^{[n]} : \alpha P \in \mathcal{S}_{\mathcal{M}}^{[n]} \}. \quad (2.4)$$

Indeed, we have the following equivalent assertions:

$$\begin{aligned} \alpha \in \mathcal{S}_{\mathcal{M}}^{[n+1]} &\iff k = 0, \dots, n \quad \alpha[P^{k+1}V_{\varphi} - V_{\varphi}\widehat{P}^{k+1}] = \mathbf{0} \quad \text{from (2.3)} \\ &\iff \alpha \in \mathcal{S}_{\mathcal{M}}^{[n]} \text{ and } \alpha P P^k V_{\varphi} - \alpha V_{\varphi} \widehat{P} \widehat{P}^k = \mathbf{0}, \quad k = 1, \dots, n \\ &\iff \alpha \in \mathcal{S}_{\mathcal{M}}^{[n]} \text{ and } \alpha P P^k V_{\varphi} - \alpha P V_{\varphi} \widehat{P}^k = \mathbf{0}, \quad k = 1, \dots, n \\ &\quad \text{since } \alpha \in \mathcal{S}_{\mathcal{M}}^{[1]} \\ &\iff \alpha \in \mathcal{S}_{\mathcal{M}}^{[n]} \text{ and } \alpha P [P^k V_{\varphi} - V_{\varphi} \widehat{P}^k] = \mathbf{0}, \quad k = 1, \dots, n \\ &\iff \alpha \in \mathcal{S}_{\mathcal{M}}^{[n]} \text{ and } \alpha P \in \mathcal{S}_{\mathcal{M}}^{[n]}. \end{aligned}$$

Next, it follows from (2.4) that

$$\mathcal{S}_{\mathcal{M}}^{[n+1]} = \mathcal{S}_{\mathcal{M}}^{[n]} \iff \mathcal{S}_{\mathcal{M}}^{[n]} P := \{ \alpha P : \alpha \in \mathcal{S}_{\mathcal{M}}^{[n]} \} \subseteq \mathcal{S}_{\mathcal{M}}^{[n]}.$$

When the inclusion in the right-hand side of the above statement holds, $\mathcal{S}_{\mathcal{M}}^{[n]}$ is said to be invariant under the matrix P . When $\mathcal{S}_{\mathcal{M}}^{[n+1]} = \mathcal{S}_{\mathcal{M}}^{[n]}$, it is easily seen from (2.4) that $\mathcal{S}_{\mathcal{M}}^{[n]} = \mathcal{S}_{\mathcal{M}}^{[n+j]}, j \geq 0$. We deduce from Representation (2.2) of $\mathcal{S}_{\mathcal{M}}(\widehat{P})$ that $\mathcal{S}_{\mathcal{M}}(\widehat{P}) = \mathcal{S}_{\mathcal{M}}^{[n+j]}, j \geq 0$.

Now, suppose the following statement to be true

$$\dim(\mathcal{S}_M^{[n+1]}) = \dim(\mathcal{S}_M^{[n]}) \implies \mathcal{S}_M^{[n+1]} = \mathcal{S}_M^{[n]}. \quad (2.5)$$

The sequence $(\dim(\mathcal{S}_M^{[n]}))_{n \geq 1}$, is nonincreasing. If there exists $n \geq 1$ such that $\dim(\mathcal{S}_M^{[n+1]}) = \dim(\mathcal{S}_M^{[n]})$, then $\mathcal{S}_M^{[n+1]} = \mathcal{S}_M^{[n]} = \mathcal{S}_M(\widehat{P})$ from (2.5) and (2.2). Consequently, the equality $\mathcal{S}_M(\widehat{P}) = \mathcal{S}_M^{[N+1]}$ follows from $\dim(\mathcal{S}_M^{[1]}) \leq N$.

The proof will be complete if we show that Assertion (2.5) is valid. Let us assume that $\dim(\mathcal{S}_M^{[n+1]}) = \dim(\mathcal{S}_M^{[n]})$. For any $n \geq 1$, the pointed polyhedral cone $\mathcal{S}_M^{[n]}$ is such that the linear subspace $\text{span}(\mathcal{S}_M^{[n]})$ generated by the vectors of $\mathcal{S}_M^{[n]}$ is equal to $\mathcal{S}_M^{[n]} - \mathcal{S}_M^{[n]}$ (see [2]). Since $\text{span}(\mathcal{S}_M^{[n+1]}) \subseteq \text{span}(\mathcal{S}_M^{[n]})$, we have

$$\text{span}(\mathcal{S}_M^{[n+1]}) = \text{span}(\mathcal{S}_M^{[n]}).$$

Suppose that there exists $\alpha \in \mathcal{S}_M^{[n]} \setminus \mathcal{S}_M^{[n+1]}$ and $\alpha \neq \mathbf{0}$. It follows from the last equality that $\alpha \in \text{span}(\mathcal{S}_M^{[n+1]})$ and $\alpha = u - v$ with $u, v \in \mathcal{S}_M^{[n+1]}$. Let us check that $\alpha P \in \mathcal{S}_M^{[n]}$: for $k = 1, \dots, n$

$$\begin{aligned} \alpha P [P^k V_\varphi - V_\varphi \widehat{P}^k] &= (u - v) [P^{k+1} V_\varphi - P V_\varphi \widehat{P}^k] \\ &= u [P^{k+1} V_\varphi - V_\varphi \widehat{P}^{k+1}] - v [P^{k+1} V_\varphi - V_\varphi \widehat{P}^{k+1}] \text{ since } u, v \in \mathcal{S}_M^{[1]}. \\ &= \mathbf{0} \end{aligned}$$

since $u, v \in \mathcal{S}_M^{[n+1]}$. Because $\alpha \in \mathcal{S}_M^{[n]}$ and $\alpha P \in \mathcal{S}_M^{[n]}$, we obtain $\alpha \in \mathcal{S}_M^{[n+1]}$ from (2.4), which is absurd. \square

Remark 2.1 *The core of the proof of the previous theorem provides a finite algorithm to compute $\mathcal{S}_M(\widehat{P})$. First, compute the extremal rays of the cone $\mathcal{S}_M^{[1]}$. Next, if $\mathcal{S}_M^{[1]}P \subseteq \mathcal{S}_M^{[1]}$ then stop with $\mathcal{S}_M(\widehat{P}) = \mathcal{S}_M^{[1]}$; otherwise compute the extremal rays of the cone $\mathcal{S}_M^{[2]}$. If $\mathcal{S}_M^{[2]}P \subseteq \mathcal{S}_M^{[2]}$ then stop with $\mathcal{S}_M(\widehat{P}) = \mathcal{S}_M^{[2]}$; otherwise compute the extremal rays of the cone $\mathcal{S}_M^{[2]}$. And so on. We see from (2.3), that a well-suited algorithm for computing the extremal rays of cones $\mathcal{S}_M^{[n]}$ is Chernikova's incremental procedure [12]. It is known from [12] that the computation of the extremal rays of a cone is exponential in dimensions involved.*

2.2 Basic properties of set $\mathcal{S}_M(\widehat{P})$

The set $\mathcal{S}_M(\widehat{P})$ is a polyhedral cone from Theorem 2.1. Let us give some basic properties of this cone.

Theorem 2.2

1. $\mathcal{S}_M(\widehat{P}) \neq \{\mathbf{0}\}$ if and only if there exists a non-trivial (polyhedral) sub-cone \mathcal{S} of $\mathcal{S}_M^{[1]}$ which is invariant under the matrix P .

2. $\mathcal{S}_{\mathcal{M}}(\hat{P})$ is the maximal convex sub-cone of $\mathcal{S}_{\mathcal{M}}^{[1]}$ which is invariant under the matrix P .
3. $\mathcal{S}_{\mathcal{M}}(\hat{P}) \neq \{\mathbf{0}\}$ iff there exists a $M \times M$ nonnegative matrix Q and a nonnegative $M \times N$ matrix U ($1 \leq M \leq N$) such that

$$UP = QU \quad \text{and} \quad U[PV_{\varphi} - V_{\varphi}\hat{P}] = \mathbf{0} \quad (2.6)$$

with $\dim \text{Cone}(U) = M$. In such a case, $\text{Cone}(U) \subseteq \mathcal{S}_{\mathcal{M}}(\hat{P})$.

4. If $\mathcal{S}_{\mathcal{M}}(\hat{P}) \neq \{\mathbf{0}\}$ then $\mathcal{S}_{\mathcal{M}}(\hat{P})$ contains one stochastic eigenvector of P corresponding to the eigenvalue 1.

Proof. 1. If $\mathcal{S}_{\mathcal{M}}(\hat{P}) \neq \{\mathbf{0}\}$ then the polyhedral cone $\mathcal{S}_{\mathcal{M}}(\hat{P})$ is included in $\mathcal{S}_{\mathcal{M}}^{[1]}$ by definition. We just have to check that $\mathcal{S}_{\mathcal{M}}(\hat{P})$ is invariant under the matrix P . We have $\mathcal{S}_{\mathcal{M}}(\hat{P}) = \mathcal{S}_{\mathcal{M}}^{[N+1]}$ from Theorem 2.1. Then $\mathcal{S}_{\mathcal{M}}^{[N+2]} = \mathcal{S}_{\mathcal{M}}^{[N+1]}$ which gives that the non-trivial cone $\mathcal{S}_{\mathcal{M}}^{[N+1]}$ is invariant under P (see the first part of the proof of Theorem 2.1).

Conversely, assume that there exists a (polyhedral) sub-cone \mathcal{S} of $\mathcal{S}_{\mathcal{M}}^{[1]}$ that is invariant under the matrix P . It follows that, for every $\alpha \in \mathcal{S}$, we have

$$\alpha P^n \in \mathcal{S}_{\mathcal{M}}^{[1]}, \quad \forall n \geq 0.$$

Therefore, we can write

$$\begin{aligned} \forall n \geq 0, \quad \alpha P^n \in \mathcal{S}_{\mathcal{M}}^{[1]} &\iff \forall n \geq 0, \quad \alpha P^{n+1}V_{\varphi} = \alpha P^n V_{\varphi} \hat{P} \quad \text{from (2.3)} \\ &\implies \forall n \geq 0, \quad \alpha P^{n+1}V_{\varphi} = \alpha V_{\varphi} \hat{P}^n \hat{P} \\ &\implies \forall n \geq 0, \quad \alpha [P^{n+1}V_{\varphi} - V_{\varphi} \hat{P}^{n+1}] = \mathbf{0} \\ &\implies \alpha \in \mathcal{S}_{\mathcal{M}}(\hat{P}) \quad \text{from (2.2)}. \end{aligned}$$

Thus, we have $\mathcal{S} \subseteq \mathcal{S}_{\mathcal{M}}(\hat{P})$.

Statement 2. is clear from the last inclusion.

3. $\mathcal{S}_{\mathcal{M}}(\hat{P})$ is a polyhedral sub-cone of \mathbb{R}_+^N , so that there exists a finite set of nonnegative vectors $U := \{u_1, \dots, u_M\}$ such that $\text{Cone}(U) = \mathcal{S}_{\mathcal{M}}(\hat{P})$, where $\text{Cone}(U)$ denotes the cone generated by the vectors u_1, \dots, u_M . Let us identify the set of row vectors U to a $M \times N$ matrix. The invariance of $\text{Cone}(U)$ under the matrix P , i.e. $\text{Cone}(U)P \subseteq \text{Cone}(U)$, shows that there exists a $M \times M$ nonnegative matrix Q such that

$$UP = QU.$$

The i th-row of matrix Q is the vector of the nonnegative coordinates of $u_i P$ with respect to the generators u_1, \dots, u_M . The equality $U[PV_{\varphi} - V_{\varphi}\hat{P}] = \mathbf{0}$ follows from the inclusion $\{u_1, \dots, u_M\} \subseteq \mathcal{S}_{\mathcal{M}}^{[1]}$.

Conversely, suppose that there exists a matrix U such that the equalities in (2.6) hold. Then, the rows of U define a polyhedral sub-cone of \mathbb{R}_+^N which is included in $\mathcal{S}_{\mathcal{M}}^{[1]}$ and is invariant under the matrix P . We obtain $\text{Cone}(U) \subseteq \mathcal{S}_{\mathcal{M}}(\hat{P})$ from Statement 2.

4. The last statement follows from the invariance of cone $\mathcal{S}_{\mathcal{M}}(\hat{P})$ under the matrix P and from [10, Lemma 3.3]. \square

Assuming the matrix P to be irreducible and $\mathcal{S}_{\mathcal{M}}(\hat{P}) \neq \{\mathbf{0}\}$, it follows from Theorem 2.2 that the stationary distribution π of P belongs to the set $\mathcal{S}_{\mathcal{M}}(\hat{P})$. Since $\pi \in \mathcal{S}_{\mathcal{M}}^{[1]}$, πV_φ is a stationary distribution of the stochastic matrix \hat{P} . In fact, we obviously have

$$\mathcal{S}_{\mathcal{M}}(\hat{P}) \neq \{\mathbf{0}\} \iff \pi V_\varphi \text{ is a stationary distribution of matrix } \hat{P}.$$

Comparison of sets $\mathcal{S}_{\mathcal{M}}(\hat{P})$ and $\mathcal{C}_{\mathcal{M}}(\hat{P})$ Let us denote the set of all $\text{distr}(X_0)$ such that $(X_n)_n$ is weakly lumpable with the matrix \hat{P} by $\mathcal{C}_{\mathcal{M}}(\hat{P})$. Let \mathcal{C}_y be the range $\mathcal{C}_{\mathcal{M}}(\hat{P})\Pi_y := \{\alpha\Pi_y : \alpha \in \mathcal{C}_{\mathcal{M}}(\hat{P})\}$ of $\mathcal{C}_{\mathcal{M}}(\hat{P})$ under the lumping projector Π_y . If $\mathcal{C}_{\mathcal{M}}(\hat{P}) \neq \{\mathbf{0}\}$, then some of the cones \mathcal{C}_y ($y \in \hat{S}$) are non-trivial and the family $(\mathcal{C}_y, y \in \hat{S})$ satisfies [10, Th 3.4, Th 3.5]

$$\mathcal{C}_y \subseteq \mathcal{C}(\hat{P}) \tag{2.7}$$

$$\mathcal{C}_y \subseteq \mathcal{C}_{\mathcal{M}}(\hat{P}) \tag{2.8}$$

$$\mathcal{C}_y \Pi_y P \Pi_x \subseteq \mathcal{C}_x \quad \text{for } x = 1, \dots, \hat{N}, \tag{2.9}$$

where

$$\mathcal{C}(\hat{P}) := \{v \geq \mathbf{0} : v\Pi_y[PV_\varphi - V_\varphi\hat{P}] = \mathbf{0}, y \in \hat{S}\}. \tag{2.10}$$

The first condition asserts that each \mathcal{C}_y is a sub-cone of the cone $\mathcal{C}(\hat{P})$. The second condition says that $\mathcal{C}_y = \mathcal{C}_{\mathcal{M}}(\hat{P})\Pi_y$ is a sub-cone of $\mathcal{C}_{\mathcal{M}}(\hat{P})$, or $\mathcal{C}_{\mathcal{M}}(\hat{P})$ is invariant under the lumping projectors Π_y ($y \in \hat{S}$). Therefore, $\mathcal{C}_{\mathcal{M}}(\hat{P})$ is the direct sum of its \hat{N} projections \mathcal{C}_y ($y \in \hat{S}$). Condition (2.9) means: the range of \mathcal{C}_y under the matrix $\Pi_y P \Pi_x$ is included in \mathcal{C}_x for every $x = 1, \dots, \hat{N}$. In particular, the cone \mathcal{C}_y is invariant under the matrix $\Pi_y P \Pi_y$ and $\mathcal{C}_{\mathcal{M}}(\hat{P})P \subseteq \mathcal{C}_{\mathcal{M}}(\hat{P})$. For each $y \in \hat{S}$, the cone \mathcal{C}_y is the maximal sub-cone of $\mathcal{C}(\hat{P})$ satisfying the three conditions (2.7)–(2.9) [10, Remark 1]. The computation of $\mathcal{C}_{\mathcal{M}}(\hat{P})$ requires to compute the extremal rays of the \hat{N} cones \mathcal{C}_y , ($y \in \hat{S}$) under the constraints (2.7)–(2.9).

If we are interested in the computation of $\mathcal{S}_{\mathcal{M}}(\hat{P})$, we only have to determine the extremal rays of the maximal sub-cone of $\mathcal{S}_{\mathcal{M}}^{[1]}$ that is invariant under the matrix P . Note the inclusion $\mathcal{C}(\hat{P}) \subseteq \mathcal{S}_{\mathcal{M}}^{[1]}$. Since $\mathcal{C}_{\mathcal{M}}(\hat{P})$ is invariant under P , we retrieve the inclusion $\mathcal{C}_{\mathcal{M}}(\hat{P}) \subseteq \mathcal{S}_{\mathcal{M}}(\hat{P})$. The major difference between the two sets $\mathcal{C}_{\mathcal{M}}(\hat{P})$ and $\mathcal{S}_{\mathcal{M}}(\hat{P})$ is that $\mathcal{S}_{\mathcal{M}}(\hat{P})$ is no more invariant under the lumping projectors $\Pi_y, y \in \hat{S}$ (*i.e.* Conditions (2.7), (2.8) may fail).

Now, let us compare the two sets for the Markov model briefly discussed in the Introduction. Specifically, we prove the claims given in Example 1.1.

Example 2.3 (Example 1.1 continued)

The Markov model discussed in Example 1.1 is based on a generalization of Cheung's software reliability model [4] provided in [11]. Cheung's model is integrated in the Cleanroom Reliability Manager whose aim is planning and certification of component-based software system reliability [16]. The basic framework of [11] is as follows. The model results in the combination of an execution model of a program with a failure model. First, the architecture of the software is represented by an HMC $(X_n)_n$ with tpm Q . $(X_n)_n$ may be thought of as the control flow graph of the program. Thus, each state i stands for a group of statements. $Q(i, j)$ is interpreted as the conditional probability for the statement group j to be executed, given that the program has just completed the execution of the statement group i . The tpm Q is here

$$Q = \begin{pmatrix} 0 & 1/7 & 2/7 & 2/7 & 1/7 & 1/7 \\ 14/70 & 0 & 19/70 & 19/70 & 14/70 & 4/70 \\ 2/8 & 2/8 & 2/8 & 0 & 1/8 & 1/8 \\ 0 & 0 & 2/8 & 4/8 & 1/8 & 1/8 \\ 7/48 & 7/48 & 14/48 & 14/48 & 0 & 6/48 \\ 5/72 & 15/72 & 20/72 & 19/72 & 13/72 & 0 \end{pmatrix}.$$

Second, the failure process is as follows. When the execution model is in state i , a failure occurs with probability p_i , thus depending on the identity of the state. For the sake of simplicity, we do not consider here that a failure may happen during a transfer of control between two groups of statements and that a failure can cause an execution break. However, a failure may affect the execution process. Indeed, state j is entered after a failure occurrence in state i with constant probability $\alpha_{i,j}$ (with $\sum_{j=1}^6 \alpha_{i,j} = 1$). The parameters are

$$\begin{aligned} p_1 &= p_2 = 1/8, & p_3 &= p_4 = 0, & p_5 &= 1/7, & p_6 &= 1/10 \\ (\alpha_{1,1} &= 1, \alpha_{1,j} = 0, j = 2, \dots, 6) & (\alpha_{2,2} &= 1/2, \alpha_{2,1} = \alpha_{2,j} = 1/10, j = 3, \dots, 6) \\ (\alpha_{3,j} &= \alpha_{4,j} = 0, j = 1, \dots, 6) & (\alpha_{5,j} &= 0, j = 1, \dots, 4, \alpha_{5,5} = 7/8, \alpha_{5,6} = 1/8) \\ (\alpha_{6,j} &= 0, j = 1, \dots, 3, \alpha_{6,4} = 1/8, \alpha_{6,5} = 1/4, \alpha_{6,6} = 5/8). \end{aligned}$$

Let us define $X^* := (X_n^*)_n$ where X_n^* is the occupied state at n . X^* is an HMC with a tpm P which is defined componentwise by $P(i, j) := Q(i, j)(1 - p_i) + p_i \alpha_{i,j}$. With the above values of parameters, P is the matrix in Example 1.1-(1.1). The stationary distribution of P is $\pi = (73, 72, 145, 145, 81, 64)/580$.

As in Example 1.1, suppose that we are interested in the expected number of visits to $\{1, 2, 3\}$ or $\{4, 5, 6\}$. Then, we need the vectors $\alpha P^n V_\varphi$ $n \geq 1$, where V_φ is as in Example 1.1 and the stochastic vector α corresponds to $\text{distr}(X_0)$.

From Theorem 2.2, (π, P) must be weakly lumpable for the state vector to obtain $\mathcal{S}_M(\hat{P}) \neq \{\mathbf{0}\}$. Put $\hat{P} := U_\pi P V_\varphi$ (see (1.2)). The stationary distribution of \hat{P} is $\hat{\pi} =$

$(1/2, 1/2)$. (π, P) is weakly lumpable for the state vector since $\hat{\pi} = \pi V_\varphi$ and

$$\forall n \geq 0, \quad \pi P^n V_\varphi = \pi V_\varphi = \hat{\pi} = \hat{\pi} \hat{P}^n.$$

- First, let us compute the set $\mathcal{S}_\mathcal{M}(\hat{P})$ along with Remark 1. Form the matrix involved in the definition of the set $\mathcal{S}_\mathcal{M}^{[1]}$ (see (2.3))

$$PV_\varphi - V_\varphi \hat{P} = \frac{1}{8} \begin{pmatrix} -1 & -1 & 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 & -1 & -1 \end{pmatrix}^\top.$$

Since this matrix is non trivial, P is not strongly lumpable according to Proposition 2.1. Now, $\mathcal{S}_\mathcal{M}^{[1]}$ is the set of all nonnegative vectors α such that $\alpha(1, 1, -1, 1, -1, -1)^\top = 0$. It is easily checked that

$$\begin{aligned} \mathcal{S}_\mathcal{M}^{[1]} &= \text{Cone}(g_i, i = 1, \dots, 9) \\ \text{with } g_1 &= (1, 0, 1, 0, 0, 0), \quad g_2 = (0, 1, 1, 0, 0, 0), \quad g_3 = (0, 0, 1, 1, 0, 0), \\ g_4 &= (1, 0, 0, 0, 1, 0), \quad g_5 = (0, 1, 0, 0, 1, 0), \quad g_6 = (0, 0, 0, 1, 1, 0), \\ g_7 &= (0, 0, 0, 1, 0, 1), \quad g_8 = (0, 1, 0, 0, 0, 1), \quad g_9 = (1, 0, 0, 0, 0, 1). \end{aligned}$$

For $i = 1, \dots, n$, the generator g_i of the cone $\mathcal{S}_\mathcal{M}^{[1]}$ is such that the nonnegative vector $g_i P$ is orthogonal to vector $(1, 1, -1, 1, -1, -1)^\top$. That is, the vectors $g_i P$ $i = 1, \dots, 9$ are in $\mathcal{S}_\mathcal{M}^{[1]}$. Consequently, $\mathcal{S}_\mathcal{M}^{[1]}$ is invariant under P and is obviously the maximal polyhedral cone included in $\mathcal{S}_\mathcal{M}^{[1]}$. We deduce from Theorem 2.2

$$\mathcal{S}_\mathcal{M}(\hat{P}) = \mathcal{S}_\mathcal{M}^{[1]}.$$

In connection with the discussion on the sets $\mathcal{S}_\mathcal{M}(\hat{P})$ and $\mathcal{C}_\mathcal{M}(\hat{P})$, note that $\pi \Pi_1 \in \mathcal{S}_\mathcal{M}(\hat{P})$. Then $\pi \Pi_1 P \in \mathcal{S}_\mathcal{M}(\hat{P})$ from Theorem 2.2. But, it is easily checked that the projection $\pi \Pi_1 P \Pi_1$ of the vector $\pi \Pi_1 P$ is not in $\mathcal{S}_\mathcal{M}(\hat{P})$.

- Second, let us consider the set $\mathcal{C}_\mathcal{M}(\hat{P})$. Since P is irreducible, we know from [9] that $\mathcal{C}_\mathcal{M}(\hat{P}) \neq \{\mathbf{0}\}$ implies that $\pi \in \mathcal{C}_\mathcal{M}(\hat{P})$. It follows from (2.10)

$$\begin{aligned} \mathcal{C}(\hat{P}) &= \{v \geq \mathbf{0} : v(1) + v(2) - v(3) = 0; v(4) - v(5) - v(6) = 0\} \quad (2.11) \\ &= \text{Cone}(g_1, g_2, g_6, g_7) \\ \text{and } \pi &= (73g_1 + 72g_2 + 81g_6 + 64g_7)/580 \in \mathcal{C}(\hat{P}). \end{aligned}$$

We must have $\pi \Pi_1 P \Pi_1 \in \mathcal{C}_1 \subseteq \mathcal{C}(\hat{P})$ from (2.7) and (2.9). But $\pi \Pi_1 P \Pi_1 \propto (471, 399, 580, 0, 0, 0) \notin \mathcal{C}(\hat{P})$ from (2.11). Hence, Condition (2.9) does not hold and $\mathcal{C}_\mathcal{M}(\hat{P}) = \emptyset$.

3 Connection between lumpability and Chapman–Kolmogorov’s condition

The first order transition probabilities of process $(\varphi(X_n))_n$, with $\text{distr}(X_0) = \alpha$, satisfy the *Chapman-Kolmogorov condition*, if there exists a $\hat{N} \times \hat{N}$ stochastic matrix \hat{P} such that

$$\forall y \in \hat{S} \text{ such that } \alpha \Pi_y \neq \mathbf{0}, \forall n \geq 1, \quad \alpha^{(y)} P^n V_\varphi = \hat{e}_y \hat{P}^n \quad (3.1)$$

where \hat{e}_y is the y -th vector of the canonical basis of $\mathbb{R}^{\hat{N}}$. We prove the following result.

Proposition 3.1 *Let α be a stochastic vector corresponding to $\text{distr}(X_0)$. The first order transition probabilities of the lumped process $\varphi(X_n)_n$ satisfy the Chapman-Kolmogorov condition with the matrix \hat{P} iff*

$$\forall y \in \hat{S}, \quad \alpha \Pi_y \in \mathcal{S}_{\mathcal{M}}(\hat{P}). \quad (3.2)$$

In other words, the probability distributions $\text{distr}(X_0)$ for which the first order transition probabilities of $\varphi(X_n)_n$ satisfy the Chapman-Kolmogorov equation with the matrix \hat{P} , are the stochastic vectors contained in $\mathcal{C}(\hat{P}) \cap \mathcal{S}_{\mathcal{M}}(\hat{P})$ ($\mathcal{C}(\hat{P})$ is defined by (2.10)).

Proof.

$$\begin{aligned} \alpha \text{ satisfies (3.1)} &\iff \forall y \in \hat{S}, \forall n \geq 1, \quad \alpha \Pi_y P^n V_\varphi = \alpha \Pi_y \mathbf{1}^\top \hat{e}_y \hat{P}^n \\ &\iff \forall y \in \hat{S}, \forall n \geq 1, \quad \alpha \Pi_y [P^n V_\varphi - \mathbf{1}^\top \hat{e}_y \hat{P}^n] = \mathbf{0} \\ &\iff \forall y \in \hat{S}, \forall n \geq 1, \quad \alpha \Pi_y [P^n V_\varphi - V_\varphi \hat{P}^n] = \mathbf{0} \\ &\quad \text{since } \alpha \Pi_y \mathbf{1}^\top \hat{e}_y = \alpha \Pi_y V_\varphi \\ &\iff \forall y \in \hat{S}, \quad \alpha \Pi_y \in \mathcal{S}_{\mathcal{M}}(\hat{P}). \end{aligned}$$

□

In the case of Example 2.3, we have $\mathcal{C}(\hat{P}) \cap \mathcal{S}_{\mathcal{M}}(\hat{P}) = \mathcal{C}(\hat{P}) = \text{Cone}(g_1, g_2, g_6, g_7)$. Thus, for any stochastic vector in $\text{Cone}(g_1, g_2, g_6, g_7)$ as $\text{distr}(X_0)$, the first order transition probabilities of $(\varphi(X_n))_n$ satisfy the Chapman-Kolmogorov condition with \hat{P} given in (1.2).

When (3.2) holds for some α and $\alpha V_\varphi > \mathbf{0}$, $(X_n)_n$ is weakly lumpable for the state vector with \hat{P} given by

$$\hat{P} = U_\alpha P V_\varphi.$$

4 Proportional dynamics for the state probabilities vector

Some authors (e.g. Tsertsvadze [21], Nicola [13] [14], Buchholz [3]) have looked for conditions under which the transient characteristics of $(X_n)_n$ are easily deduced from those of

$(\varphi(X_n))_n$. A first instance of such a condition is related with what we call a *Rogers-Pitman (R-P) matrix* P . That is, P satisfies

$$\Lambda P = \widehat{P} \Lambda, \quad (4.1)$$

where Λ is a stochastic φ -block diagonal matrix and \widehat{P} is some $\widehat{N} \times \widehat{N}$ stochastic matrix. In such a case, $(\varphi(X_n))_n$ is an HMC with tpm \widehat{P} , at least for any $\widehat{e}_y \Lambda$ as $\text{distr}(X_0)$. For such a R-P matrix, the transient (and stationary) state vector for $(X_n)_n$ may be derived from that of the lumped model. Indeed, it follows from (4.1) that

$$\forall n \geq 1, \quad \Lambda P^n = \Lambda P P^{n-1} = \widehat{P} \Lambda P^{n-1} = \widehat{P}^n \Lambda. \quad (4.2)$$

Thus, with any $\widehat{e}_y \Lambda$ as $\text{distr}(X_0)$, the one-dimensional distributions of the HMC $(X_n)_n$ are obtained from a computation with a $\widehat{N} \times \widehat{N}$ matrix. Such a relation is known [6], [19], [3] when P is a Schweitzer's matrix (that is, $\Lambda = U_1$).

In the context of linear dynamical systems, a condition was identified in [21] for lumping with respect to the state probabilities vector. In this work, note that α and \widehat{P} are not constrained to be nonnegative. A similar concept is developed in [13] for continuous time positive linear systems and is referred as to the *proportional dynamics property* for the transient measure concerned with, here the state probabilities vector. In our discrete time context, this property can be reformulated as follows.

Definition 4.1 *Let α be a stochastic vector on S . The Markov chains (α, P) and $(\alpha V_\varphi, \widehat{P})$ have proportional dynamics for the state probabilities vector, if there exists a stochastic φ -block diagonal matrix U such that*

$$\forall n \geq 1, \quad \alpha P^n = \alpha V_\varphi \widehat{P}^n U. \quad (4.3)$$

Such a condition has to be interpreted as follows: for α as $\text{distr}(X_0)$, for every $n \geq 1$

$$\forall y \in \widehat{S}, \forall x \in \varphi^{-1}(y), \quad \mathbb{P}_\alpha\{X_n = x\} = \mathbb{P}_\alpha\{\varphi(X_n) = y\} U(y, x).$$

Remark 4.1 *In contrast to [13], Relation (4.3) is not required for $n = 0$. It will be shown that the proportional dynamics property with α satisfying $\alpha = \alpha V_\varphi U$, is equivalent to P is a R-P matrix.*

Note that for a R-P matrix P , we have from (4.2)

$$\Lambda P^n = \widehat{P}^n \Lambda = \Lambda V_\varphi \widehat{P}^n \Lambda$$

since $\Lambda V_\varphi = \widehat{I}$. Consequently, we have the proportional dynamics property with any stochastic vector $\alpha \in \text{Cone}(\Lambda)$ as $\text{distr}(X_0)$. If Relation (4.3) holds for some stochastic vector α , then

$$\forall n \geq 1, \quad \alpha P^n V_\varphi = \alpha V_\varphi \widehat{P}^n U V_\varphi = \alpha V_\varphi \widehat{P}^n$$

and $\alpha \in \mathcal{S}_{\mathcal{M}}(\widehat{P})$. Hence, the proportional dynamics property implies that (α, P) is weakly lumpable for the state vector. The converse is generally false (see Example 4.1).

For any stochastic φ -block diagonal matrix U , let us consider as in [13], the set S_U of the stochastic vectors contained in the polyhedral cone $\mathcal{C}_U := \{\alpha \geq \mathbf{0} : \alpha V_\varphi U = \alpha\}$. S_U is the solution set of Equation (4.3) for $n = 0$. Note that we have $\widehat{e}_y U \in \mathcal{C}_U$ for every $y \in \widehat{S}$ since $UV_\varphi = \widehat{I}$. That is, we have $\text{Cone}(U) \subseteq \mathcal{C}_U$. In fact, $\mathcal{C}_U = \text{Cone}(U)$ since we can write for $\alpha \in \mathcal{C}_U$

$$\alpha = \sum_{y \in \widehat{S}} \alpha \Pi_y = \sum_{y \in \widehat{S}} (\alpha \Pi_y \mathbf{1}^\top) \widehat{e}_y U \in \text{Cone}(U).$$

Assuming in addition that (4.3) holds, we obtain from (4.3) with $n = 1$

$$UP = UV_\varphi \widehat{P}U = \widehat{P}U \quad (\text{since } UV_\varphi = \widehat{I}).$$

That is, P is a R-P matrix with $\Lambda = U$. Hence, the matrix P has the proportional dynamics property for every stochastic vector $\alpha \in \text{Cone}(U)$, where U is a stochastic φ -block diagonal matrix, is equivalent to P is a R-P matrix with U as associated matrix.

Finally, if the Markov chains (α, P) and $(\alpha V_\varphi, \widehat{P})$ have proportional dynamics for the state vector with every α , then P is a R-P matrix which is strongly lumpable. Such a matrix is called a strictly lumpable matrix in [3]. The converse is false in general as shown by the following example.

Example 4.1

Let us consider the tpm P below with a lumping map φ defined from $S = \{1, 2, 3, 4\}$ into $\{1, 2\}$ by $\varphi(1) = \varphi(2) = 1$ and $\varphi(3) = \varphi(4) = 2$:

$$P = \left(\begin{array}{cc|cc} 1/6 & 1/3 & 1/3 & 1/6 \\ 1/3 & 1/6 & 1/6 & 1/3 \\ \hline 1/3 & 1/6 & 1/4 & 1/4 \\ 1/6 & 1/3 & 1/4 & 1/4 \end{array} \right) \quad V_\varphi = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}.$$

P is a R-P matrix since $\Lambda P = \widehat{P}\Lambda$ with

$$\widehat{P} = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} \quad \Lambda = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 \end{pmatrix}.$$

Note that P is also strongly lumpable since $PV_\varphi = V_\varphi \widehat{P}$. We have the proportional dynamics property for every stochastic vector α in $\text{Cone}(\Lambda)$.

For the vector $e_1 := (1, 0, 0, 0) \notin \text{Conv}(\Lambda)$, we check that the proportional dynamics property fails to hold for any stochastic φ -block diagonal matrix U . We have

$$\forall n \geq 1, \quad e_1 V_\varphi \widehat{P}^n = (1/2 \quad 1/2).$$

Let us use

$$U = \begin{pmatrix} \alpha_1 & 1 - \alpha_1 & 0 & 0 \\ 0 & 0 & \alpha_2 & 1 - \alpha_2 \end{pmatrix} \quad \alpha_1, \alpha_2 \in [0, 1],$$

as generic matrix U . Then, the vector $e_1 V_\varphi \hat{P}^n U$ is $(1/2)(\alpha_1, 1 - \alpha_1, \alpha_2, 1 - \alpha_2)$ for every $n \geq 1$. If U is a solution of $e_1 P = e_1 V_\varphi \hat{P} U$, then $\alpha_1 = 1/3$ and $\alpha_2 = 2/3$. But, we have $e_1 P^2 = (1/72)(20, 16, 17, 19) \neq e_1 V_\varphi \hat{P}^2 U = (1/6, 1/3, 1/3, 1/6)$. Thus, the proportional dynamics property does not hold from (4.3). However, $e_1 \in \mathcal{S}_M(\hat{P})$ since the matrix P is strongly lumpable into the matrix \hat{P} .

We do not go into further details since the property of proportional dynamics is more stringent than weak lumpability. We refer to [3] for the practical use of the proportional dynamics property. We conclude this section by proving the following result.

Proposition 4.1 *For every stochastic vector α , the Markov chains (α, P) and $(\alpha V_\varphi, \hat{P})$ have proportional dynamics for the state vector with the φ -block diagonal matrix U iff*

$$P = V_\varphi \hat{P} U.$$

Proof. Assume that (α, P) and $(\alpha V_\varphi, \hat{P})$ have proportional dynamics for every initial distribution α . We have from Relation (4.3) for $n = 1$

$$\begin{aligned} \forall \alpha \geq \mathbf{0}, \quad \alpha[P - V_\varphi \hat{P} U] = \mathbf{0} &\implies \forall \alpha \in \mathbb{R}^N, \quad \alpha[P - V_\varphi \hat{P} U] = \mathbf{0} \\ &\implies P = V_\varphi \hat{P} U. \end{aligned}$$

Conversely, if $P = V_\varphi \hat{P} U$ then $P^n = (V_\varphi \hat{P} U)^n = V_\varphi \hat{P}^n U$ since $U V_\varphi = \hat{I}$. □

5 Conclusion

In this paper, we introduce a concept of lumpability with respect to the state vector of a discrete-time homogeneous finite Markov chain. The merit of this form of lumpability is twofold: it simplifies the computation of the state vector of a Markov model, and it may be applied when standard lumpability criteria fail. Finally, almost all results presented in this paper have their continuous-time counterpart using the uniformization technique (e.g. see [8, 7]).

References

- [1] AOKI, M. (1968). Control of Large-Scale Dynamic systems by Aggregation. IEEE Trans. Automat. Control **13**, 246–253.

- [2] BERMAN, A. AND PLEMMONS, R. J. (1979). *Nonnegative Matrices in the Mathematical Sciences*. Academic Press.
- [3] BUCHHOLZ, P. (1994). Exact and ordinary lumpability in finite Markov chain. *J. Appl. Probab.* **31**, 59–75.
- [4] CHEUNG R. C. (1980). A user-oriented software reliability model. *IEEE Trans. Software Eng.* **6**, 118–125.
- [5] COXSON, P. G. (1984). Lumpability and Observability of Linear Systems. *J. Math. Anal. Appl.* **99**, 435–446.
- [6] DELEBECQUE, F., QUADRAT, J. P. AND KOKOTOVIC P. V. (1984). A unified view of aggregation and coherency in networks and Markov chains. *Int. J. Control* **40**, 939–952.
- [7] GROSS, D. AND MILLER, D. R. (1984). The randomization technique as a modeling tool and solution procedure for transient Markov processes. *Oper. Res.* **32**, 343–361.
- [8] JENSEN, A. (1953). Markov chains as an aid in the study of Markov processes. *Skand. Aktuarietidsshr.* **36**, 87–91.
- [9] KEMENY, J. G. AND SNELL, J. L. (1976). *Finite Markov chains*. Springer-Verlag.
- [10] LEDOUX, J. (1997). A geometric invariant in weak lumpability of finite Markov chains. *J. Appl. Probab.* **34**, 847–858.
- [11] LEDOUX J. AND RUBINO G. (1997). Simple formulae for counting processes in reliability models. *Adv. in Appl. Probab.* **29**, 1018–1038.
- [12] MATHEISS, T. H. AND RUBIN, D. S. (1980). A survey and comparison of methods for finding all vertices of convex polyhedral sets. *Math. Oper. Res.* **5**, 167–185.
- [13] NICOLA, V. F. (1989). Lumping in Markov reward processes. Res. Rept 14719, IBM.
- [14] NICOLA, V. F. (1995). Lumping in Markov reward processes. In *Computation with Markov chains*, Ed. W.J. Stewart. Kluwer, pp 663–666.
- [15] ROGERS, L. C. G. AND PITMAN, J. W. (1981). Markov functions. *Ann. Probab.* **9**, 573–582.
- [16] POORE, J. H., MILLS H. D. AND MUTCHLER D. (1993). Planning and certifying software system reliability. *IEEE Software* **10**, 89–99.
- [17] RUBINO, G. AND SERICOLA, B. (1991). A finite characterization of weak lumpable Markov processes. Part I: The discrete time case. *Stochastic Process. Appl.* **38**, 195–204.

- [18] RUBINO, G. AND SERICOLA, B. (1990). Closed-form solution for the distribution of the total time spent in a subset of states of a homogeneous Markov process during a finite observation period. *J. Appl. Probab.* **27**, 713–719.
- [19] SCHWEITZER, P. J. (1984). Aggregation methods for large Markov chains. In *Mathematical Computer Performance and Reliability*, Eds Iazeolla and others. North Holland.
- [20] SUMITA, U. AND RIEDERS, M. (1989). Lumpability and time reversibility in the Aggregation-Disaggregation method for large Markov chains. *Stoch. Models* **5**, 63–81.
- [21] TSERTSVADZE, G. N. (1989). Aggregation of systems described by Markov chains. *Automation and Remote Control* **5**, 628–636.